2К

УДК 004.93’12

Г.А. Кобзарь
Харьковский национальный университет радиоэлектроники, г. Харьков, Украина
glebkobzar@yahoo.com

Модель межмасштабного пространства кривизны для представления формы геометрических объектов

Предлагается модель межмасштабного пространства кривизны для представления формы контуров геометрических объектов, которая в отличие от известных моделей масштабно-пространственного представления отображает выпуклые признаки формы в масштабном пространстве за счет введения межмасштабного проецирования, что позволяет значительно расширить класс объектов, распознаваемых с помощью масштабно-пространственных представлений.

Введение

Разработка информативного представления формы (дескриптора формы) геометрических объектов является одной из основных задач при построении систем распознавания геометрических объектов по форме. При возникновении задач представления и распознавания формы геометрических объектов были выдвинуты следующие необходимые требования к дескрипторам формы: инвариантность к перемещению, масштабированию и повороту. В рамках разработки стандарта «Интерфейс представления мультимедийных данных» группой экспертов MPEG (Moving Picture's Experts Group), больше известного под названием MPEG-7, были выдвинуты новые критерии оценки представлений формы: инвариантность, устойчивость, компактность, вычислительная простота формирования, возможность и вычислительная сложность сравнения.

Представление формы геометрических объектов в виде масштабного пространства кривизны (МПК) – дескриптор CSS – является одним из четырех успешно прошедших официальные испытания в рамках разработки стандарта MPEG-7. Именно этот дескриптор наиболее полно отображает локальные признаки формы. Проблемой дескриптора в плане представления локальных признаков является лишь отсутствие возможности отражения выпуклых признаков формы, что, собственно, и является основной причиной исключительно локальной ориентации этого дескриптора.

В данной статье анализируется традиционная модель масштабно-пространственного представления формы геометрических объектов и ее модификация на основе проецирования на окружность с последующим формированием дуального набора признаков, позволяющая в некоторых случаях добиться отображения выпуклых признаков формы, а также разработанная модель межмасштабного проецирования, позволяющая наиболее полно отображать все выпуклые и вогнутые признаки формы геометрических объектов на одном наборе признаков.

Статья организована следующим образом. В первой главе приводится описание исходной модели масштабно-пространственного представления кривизны (МПК), построение изображений МПК и анализ достоинств и недостатков такого

представления. Во второй главе рассматривается модификация МПК на основе введения дуального изображения МПК путем проецирования на окружность и анализ этого решения проблемы традиционной модели МПК. В третьей главе описывается предлагаемая модель межмасштабного пространства кривизны, позволяющая за счет межмасштабного проецирования добиться наиболее полного отображения выпуклых признаков. В четвертой главе рассматриваются некоторые аспекты практического применения предлагаемой модели для распознавания образов геометрических объектов.

1 Модель масштабно-пространственного представления формы

Теория многомасштабного представления измеренного сигнала была предложена Виткином (1983) и одновременно Коэндериком (1984). Методология предполагает отображение сигнала на однопараметрическое семейство полученных масштабов – масштабное пространство (МП). Параметром в этом случае является так называемый масштабный параметр $\sigma \in \mathbb{R}_+$, необходимый для определения некоторого масштабного уровня ($\mathbb{R}_+$ – множество действительных чисел, а $\mathbb{R}_+ \setminus \{0\}$ – соответствующее множество без нуля). Рассмотрим, как формулируется масштабно-пространственное представление для непрерывного сигнала (в одномерном случае) [1].

Определение 3.1. Для сигнала $f:\mathbb{R} \to \mathbb{R}$ масштабно-пространственное представление $L:\mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ определено так, что отображение сигнала на нулевом уровне эквивалентно исходному сигналу, а отображение на более грубых масштабах выражается конволюцией исходного сигнала гауссовыми ядрами возрастающей ширины.

$$L(x;0) = f(x),\; L(x;\sigma) = g(x;\sigma) * f(x), \tag{1}$$

где $L(x;0)$ – отображение сигнала на нулевом уровне;

$L(x;\sigma)$ – отображение сигнала на более грубых масштабах σ.

В форме неявных интегралов результат операции конволюции «*» записывается в виде:

$$L(x;\sigma) = \int_{\xi=-\infty}^{+\infty} g(\xi;\sigma) f(x-\xi)\, d\xi,$$

где $g:\mathbb{R} + \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}$ является Гауссовым ядром (одномерным) [1].

Причем доказано, что именно Гауссово МП – единственное непрерывное линейное МП, удовлетворяющее всем необходимым требованиям: линейность (в отношении умножения), инвариантность к сдвигу, масштабная инвариантность, свойство полугруппы, свойство положительности.

Представляя контур замкнутой области некоторого объекта в параметрическом виде и учитывая, что контур замкнутой области является замкнутой кривой, можно представить кривую Γ с помощью параметра, нормализующего длину дуги

$$\Gamma = \{(x(u), y(u) \mid u \in [0,1]\}. \tag{2}$$

Для формулировки масштабно-пространственного представления кривизны контура замкнутой области, представленного кривой в параметрической форме, вводят понятие развития кривой, которая достигается вычислением кривизны контура на различных масштабах (сглаженной Гауссовой функцией с σ, соответствующими масштабу).

$$X(u,\sigma)=x(u)*g(u,\sigma), \qquad Y(u,\sigma)=y(u)*g(u,\sigma).$$

Согласно свойствам конволюции, производные каждого компонента могут быть легко вычислены:

$$X_u(u,\sigma)=x(u)*g_u(u,\sigma), \qquad X_{uu}(u,\sigma)=x(u)*g_{uu}(u,\sigma). \tag{3}$$

Точно так же определяются $Y(u,\sigma)_u$ и $Y(u,\sigma)_{uu}$. Поскольку точные формулы для $g_u(u,\sigma)$ и $g_{uu}(u,\sigma)$ известны, кривизну развивающейся кривой можно легко вычислить [2]:

$$k(u,\sigma)=\frac{X_u(u,\sigma)Y_{uu}(u,\sigma)-X_{uu}(u,\sigma)Y_u(u,\sigma)}{(X_u(u,\sigma)^2+Y_u(u,\sigma)^2)^{3/2}}. \tag{4}$$

При увеличении σ изменяется форма Γ_σ. Этот процесс формирования упорядоченной последовательности кривых называется развитием Γ. Рассматривая Γ, не трудно определить положение точек перехода через ноль для Γ_σ, начиная с $\sigma=1$ и увеличивая на некоторое $\Delta\sigma$ на каждом уровне масштаба (рис. 1).

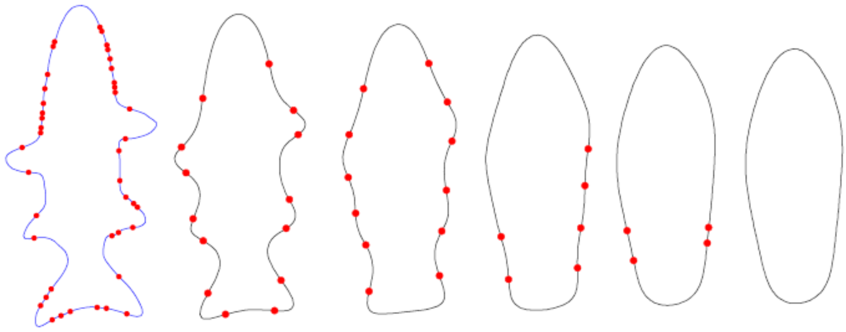
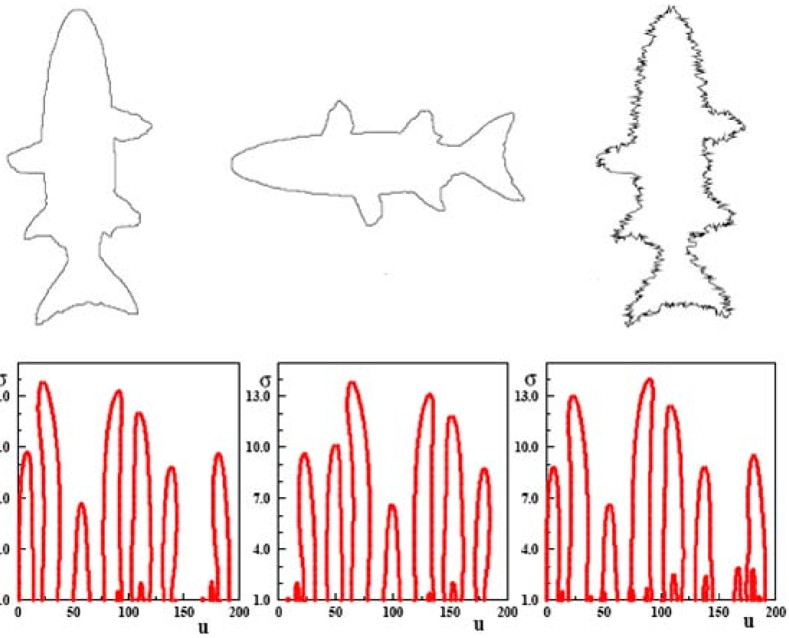


Рисунок 1 – Сглаживание кривой при развитии и уменьшение количества точек перехода через ноль при $\sigma=1,4,7,10,12,14$

Если определить положения точек перехода кривизны через ноль каждой Γ_σ при развитии, можно отобразить результирующие точки на плоскости (u,σ), где u – нормированная длина дуги, а σ – ширина Гауссова ядра [3]. Результат можно отобразить в виде бинарного изображения – изображения масштабного пространства кривизны (МПК) кривой (рис. 3, 4). Пересечение каждой горизонтальной линии контурами МПК отмечает положение точек перехода через ноль соответствующей кривой Γ_σ.

Очевидно, что изображение МПК нормированной кривой является инвариантным к аффинным искажениям и устойчиво к помехам. Поворот объекта обычно вызывает циклический сдвиг МПК (рис. 2а, 2б). Точно такое же влияние имеет изменение начальной точки контура. Благодаря нормализации масштабирование не влияет на вид МПК, а на рис. 2в видно, что помехи производят появление небольших дуг на низких уровнях σ, но никак не влияют на основные дуги.

Эти качества МПК используют для эффективного представления и распознавания формы геометрических объектов. Необходимо отметить, что для идентификации формы не требуется все изображение МПК, а только максимумы дуг. Набор максимумов в виде пар и составляет дескриптор CSS (Curvature Scale-Space – масштабного пространства кривизны, далее – МПК), выбранный в качестве одного из основных дескрипторов формы в рамках стандарта MPEG-7.



а б в

Рисунок 2 – Кривые контуров и соответствующие изображения МПК:
а – исходный контур, б – влияние поворота, в – влияние помех

Однако, несмотря на все достоинства, этот подход обладает одним недостатком – отсутствие возможности отражения даже сильно выраженных выпуклых признаков формы, т.к. выпуклые сегменты не имеют точек перехода кривизны через ноль. Очевидно, что введение выпуклых признаков в дескриптор МПК либо модификация процесса формирования дескриптора позволит значительно расширить класс распознаваемых объектов.

2 Дуальная модель масштабно-пространственного представления формы

Попытка введения выпуклых признаков в дескриптор МПК была сделана Копфом в 2005 с помощью проецирования кривых Γ_σ на описывающую окружность [4].

Для проецирования кривых Γ_σ на окружность, кривая вписывается в окружность радиуса R и каждая точка кривой проецируется на ближайшую точку окружности P. Пример проецирования на окружность кривой с сильно выраженными выпуклыми признаками (углами) представлена на рис. 3. Выпуклые сегменты исходной кривой в результате проецирования становятся вогнутыми сегментами проецированной кривой. Вычисление проецированной кривой достаточно просто.

Точки $(x(u), y(u))$ исходной замкнутой кривой проецируются в точки с другими координатами $(x'(u), y'(u))$. Координаты центра описывающей окружности (M_x, M_y) с радиусом R вычисляются как средние координат точек кривой.

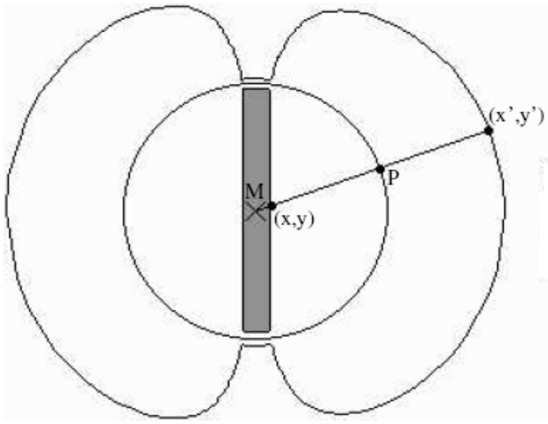


Рисунок 3 – Проецирование кривой на окружность

Непосредственно проецирование выполняется по следующим формулам:

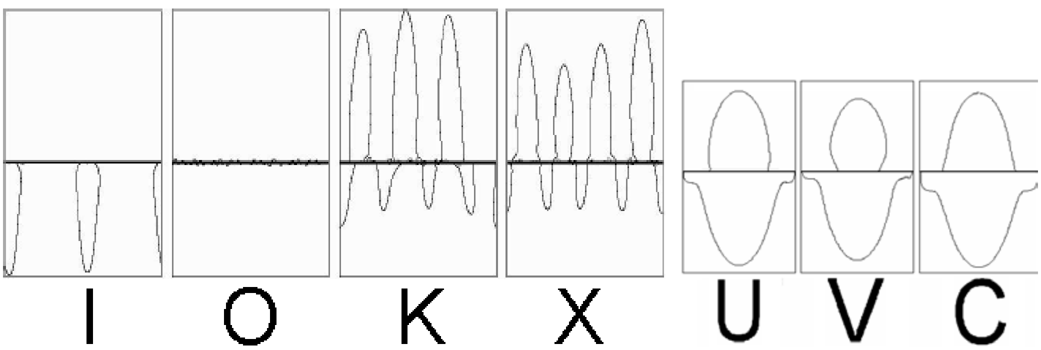
$$D_{x(u),y(u)} = \sqrt{(M_x - x(u))^2 + (M_y - y(u))^2}\,, \tag{5}$$

$$x'(u) = (x(u) - M_x)\cdot\frac{2\cdot R - D_{x(u),y(u)}}{D_{x(u),y(u)}} + M_x\,, \tag{6}$$

$$y'(u) = (y(u) - M_y)\cdot\frac{2\cdot R - D_{x(u),y(u)}}{D_{x(u),y(u)}} + M_y\,, \tag{7}$$

где $D_{x(u),y(u)}$ – расстояние между центром (M_x, M_y) и точками $(x(u), y(u))$ кривой.

В дискретном случае, если точка совпадает по координатам с центром, координаты ее проекции определяются как средние между координатами проекций двух соседних точек.



а б

Рисунок 4 – Дуальное МПК выпуклых и вогнутых символов текста:
а – удачное дуальное МПК, б – недостаток дуального МПК (дуги центрированы, что не влияет на представление, но более наглядно)

Построение МПК спроецированной кривой не отличается от построения МПК обычной кривой кроме знака. Таким образом, в результате построения МПК исходной кривой и ее проекции получают дуальное изображение МПК (рис. 4а).

Из рис. 4а видно, что такое дуальное представление МПК более полно отображает признаки формы геометрических объектов по сравнению с исходным МПК, так как в негативной части дуального МПК отражаются наиболее выраженные выпуклые признаки формы, тогда как в исходном МПК выпуклые признаки вообще отсутствуют. Однако также можно заметить, что это представление является избыточным, так как вогнутые сегменты исходной кривой отображаются в вогнутые на спроецированной кривой.

Кроме того, такое отображение не способно в достаточной мере различить степень выгнутости выпуклых признаков. Это можно легко показать на примере символов текста, для распознавания которых, как ни странно, применение дуального МПК рекомендовано (рис. 4б). На приведенном на рисунке примере проецирование контуров символов на окружность не дает ожидаемого эффекта, поскольку проекция мало отличается от исходного контура либо от других проекций. Еще одним недостатком дуального представления на основе проецирования на окружность является возникновение петель в некоторых (иногда значимых) участках контура из-за несоответствия формы кривых контура понятию округлости. При этом такие участки не отображаются на дуальном МПК.

Таким образом, проблема отображения выпуклых признаков в МПК остается открытой.

3 Межмасштабное пространство кривизны

Метод проецирования кривой на более примитивную, в смысле детализации, кривую, рассмотренный на примере окружности выше, имеет одну очень полезную особенность – представление исходной кривой таким образом, что выпуклые признаки проецируются в вогнутые. Эта особенность проявляется не во всех случаях, поскольку в качестве кривой проецирования выбрана фиксированная замкнутая кривая – окружность. Очевидно, ни одна фиксированная кривая, ввиду разнообразия форм геометрических объектов, не может дать адекватного (в смысле отображения выпуклых признаков в вогнутые) проецирования.

Предлагается масштабно-пространственное представление кривизны, построенное на основе развития кривой контура, спроецированной на кривую более высокого уровня (более округлой). Такой подход позволяет отобразить все множество выпуклых признаков на множество вогнутых признаков, исключая необходимость применения фиксированной кривой.

Как и ранее, представляем кривую контура Γ с помощью параметра, нормализующего длину дуги (2). Представление кривой на различных уровнях масштаба σ осуществляется развитием кривой путем конволюции Гауссовой функцией $g(u,\sigma)$ в соответствии с (3).

Для формулировки представления кривизны контура замкнутой области в межмасштабном пространстве введем понятие межмасштабного проецирования, при котором на каждом уровне масштаба σ кривая Γ_σ проецируется на представление этой кривой на более высоком уровне $\Gamma_{\sigma+\Delta\sigma}$ в кривую $\Gamma_\sigma^{\Delta\sigma}$ в соответствии со следующей формулой.

$$X^{\Delta\sigma}(u,\sigma) = -X(u,\sigma) + Dx(u,\sigma), \quad Y^{\Delta\sigma}(u,\sigma) = -Y(u,\sigma) + Dy(u,\sigma),$$

где $Dx(u,\sigma)$ и $Dy(u,\sigma)$ – компоненты вектора $\vec{D}=(Dx(u,\sigma),Dy(u,\sigma))$, с началом в точке $(X(u,\sigma),Y(u,\sigma))$ и концом в точке $(X(u,\sigma+\Delta\sigma),Y(u,\sigma+\Delta\sigma))$, которые могут быть легко вычислены по формулам:

$$Dx(u,\sigma)=X(u,\sigma)-X(u,\sigma+\Delta\sigma),\quad Dy(u,\sigma)=Y(u,\sigma)-Y(u,\sigma+\Delta\sigma).$$

Таким образом, спроецированная кривая может быть представлена в параметрическом виде следующим образом:

$$\begin{aligned}X^{\Delta\sigma}(u,\sigma)&=-x(u)*g(u,\sigma)+2\cdot x(u)*g(u,\sigma+\Delta\sigma),\\Y^{\Delta\sigma}(u,\sigma)&=-y(u)*g(u,\sigma)+2\cdot y(u)*g(u,\sigma+\Delta\sigma).\end{aligned}\qquad(8)$$

По аналогии с формулой (4) определяем значение кривой в межмасштабном пространстве как:

$$k(u,\sigma)=\frac{X_u^{\Delta\sigma}(u,\sigma)Y_{uu}^{\Delta\sigma}(u,\sigma)-X_{uu}^{\Delta\sigma}(u,\sigma)Y_u^{\Delta\sigma}(u,\sigma)}{(X_u^{\Delta\sigma}(u,\sigma)^2+Y_u^{\Delta\sigma}(u,\sigma)^2)^{3/2}},\qquad(9)$$

где $X_u^{\Delta\sigma},Y_u^{\Delta\sigma}$ и $X_{uu}^{\Delta\sigma},Y_{uu}^{\Delta\sigma}$ определяются по формуле (9) при подстановке $g_u(u,\sigma)$ и $g_{uu}(u,\sigma)$ вместо $g(u,\sigma)$ соответственно.

Формула (9) позволяет определить значение кривизны в каждой точке межмасштабного пространства кривизны (ММПК).

Пример проецирования кривой контура квадрата на соответствующее представление на более высоком уровне показан на рис. 5.

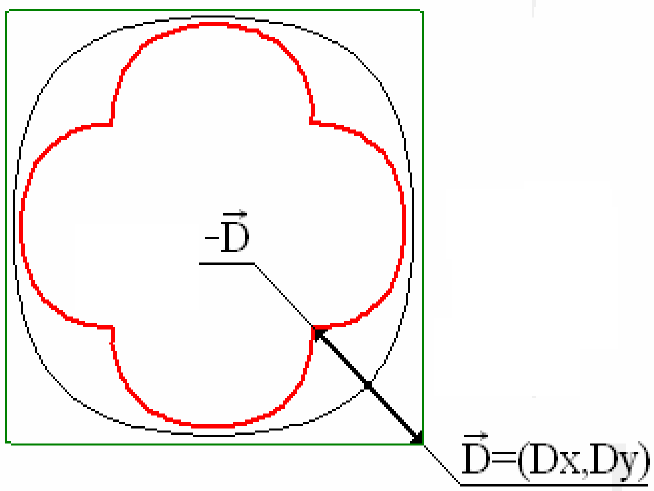


Рисунок 5 – Межмасштабное проецирование кривой контура квадрата ($\Delta\sigma=20$)

Внешний контур представляет кривую на уровне $\sigma=0$, средний – кривую на уровне $\sigma=20$, а внутренний – соответствующая спроецированная кривая. Очевидно, межмасштабное проецирование кривой Γ_σ позволяет преобразовать выпуклые признаки в вогнутые на каждом уровне масштаба.

4 Аспекты практического применения межмасштабного проецирования

Построение изображения ММПК ничем не отличается от построения МПК. При увеличении σ изменяется форма Γ_σ, а соответственно и форма $\Gamma_\sigma^{\Delta\sigma}$. Не трудно определить положение точек перехода через ноль для $\Gamma_\sigma^{\Delta\sigma}$ с помощью формулы (9), начиная с $\sigma = 1$ и увеличивая на каждом уровне масштаба. При увеличении σ и постепенном сглаживании исходной кривой сглаживается и проецированная кривая, а соответственно количество точек перехода через ноль уменьшается до тех пор, пока кривая $\Gamma_\sigma^{\Delta\sigma}$ не станет выпуклой, т.е. не останется ни одной точки перехода через ноль на проецированной кривой. Для наглядности приведем изображения МПК и ММПК одной кривой при повороте и наличии искажений (рис. 6).

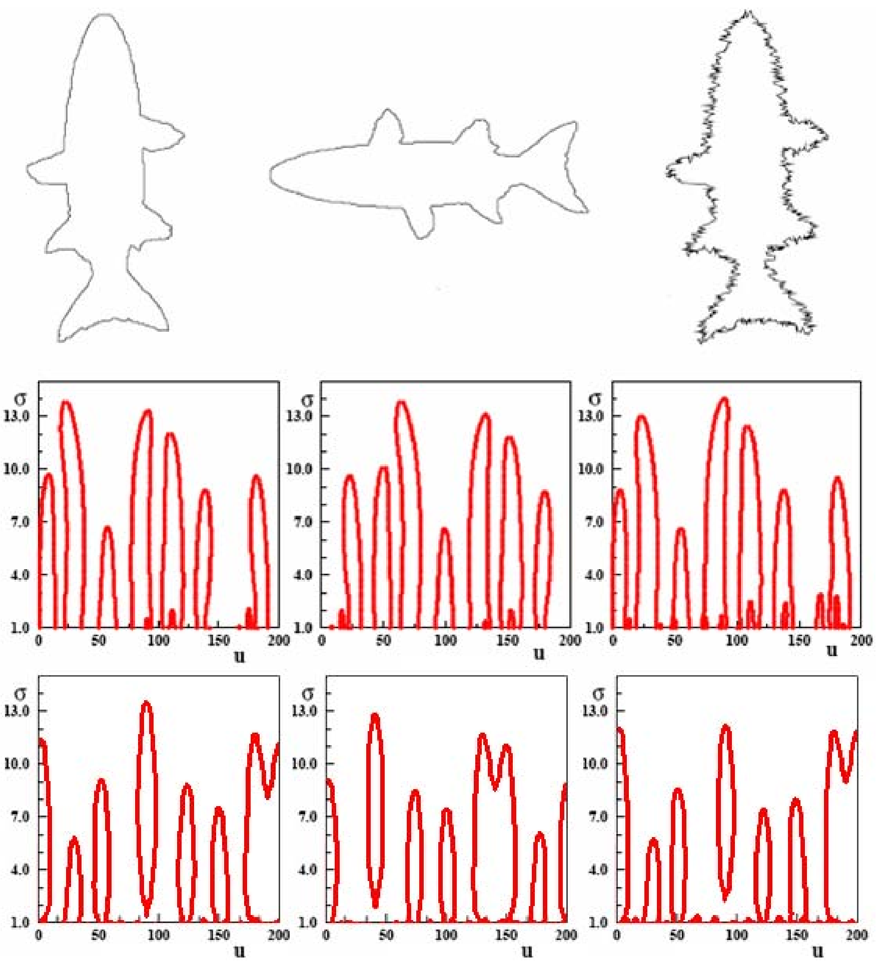


Рисунок 6 – Кривые контуров и соответствующие изображения МПК и ММПК

На рис. 6 видно, что выпуклый признак головы рыбы отображается отдельной дугой на ММПК ($u = 150$) и не имеет отображения на МПК. Кроме того, следует заметить, что дуги изображения ММПК могут иметь совершенно иную форму, нежели дуги МПК. Появляются точки, которые можно назвать точками обратного максимума (для дуг, имеющих эллиптическую форму, $u = 90$), а также точками

локальных минимумов ($u = 190$). Исследуя поведение дуг ММПК при изменении $\Delta\sigma$, можно заметить, что точки локального минимума появляются вследствие слияния двух дуг, поэтому их включение в набор признаков дескриптора может повлечь возникновение ошибок при распознавании и, следовательно, не является эффективным.

Исследуя возникновение и значимость эллиптических дуг при построении ММПК, можно заключить, что учет положения точек обратного максимума может иметь решающее значение при распознавании (рис. 7).

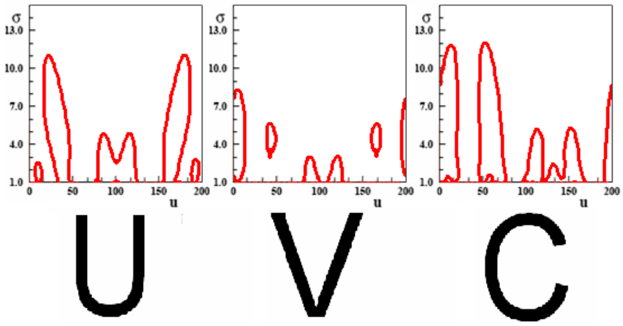


Рисунок 7 – Различие ММПК сходных по строению символов текста

Рассматривая изображения МПК (рис. 4а, верхний набор), дуального МПК (рис. 4а, верхний и нижний наборы) и ММПК (рис. 7) букв «U», «V» и «C» латинского алфавита, можно заключить, что межмасштабно-пространственное представление кривизны замкнутых кривых обеспечивает наиболее полное отображение локальных признаков формы геометрических объектов. Это создает предпосылки создания более адекватного дескриптора формы и, следовательно, более эффективных методов распознавания геометрических объектов по признакам формы.

Заключение

Анализ традиционного метода построения масштабно-пространственных представлений кривизны (МПК) контуров геометрических объектов позволил выявить причины отсутствия выпуклых признаков в таких представлениях. Анализ предлагаемого для решения этой проблемы дуального МПК выявляет избыточность и неадекватность данного представления, однако дает предпосылки возможного решения проблемы отображения выпуклых признаков путем проецирования.

Введенная формулировка межмасштабного проецирования дает возможность разработки нового масштабно-пространственного представления формы геометрических объектов, включающего и вогнутые, и выпуклые признаки формы. Представление формы геометрических объектов с помощью модели межмасштабного пространства кривизны обеспечивает наиболее полное отображение признаков формы геометрических объектов, что создает предпосылки создания более эффективных методов распознавания геометрических объектов по признакам формы, позволяя значительно расширить класс распознаваемых объектов.

Планируется проведение широких экспериментальных исследований разработанной модели на базах геометрических объектов, распознавание которых было ранее не возможно либо не адекватно с использованием масштабно-пространственных представлений из-за ограничений исходной модели, а также сравнения дескриптора, полученного на основе модели, с иными, принятыми в рамках стандарта MPEG-7.

На основе данного теоретического материала, а также опираясь на результаты экспериментальных исследований, планируется публикация также и в зарубежных научных изданиях.

Литература

1. Linderberg T. Scale Space Theory in Computer Vision // Kluwer Academic Publishers. – 1994. – 440 p.
2. Mokhtarian F.and Mackworth A.K. A theory of multi-scale curvature based shape representation for planar curves // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1986. – Vol. 8, № 1. – P. 34-43.
3. Mokhtarian F., Abbasi S., Kittler J. Efficient and Robust Retrieval by Shape Content through Curvature Scale Space // Int. Workshop on Image DataBases and Multimedia Search. – 1996. – P. 35-42.
4. Kopf S., Haenselmann T., Effelsberg W. Enhancing curvature scale space features for robust shape classification // Proceedings to IEEE International Conference on Multimedia and Expo (ICME). – 2005. – P. 478-481.

Г.А. Кобзар
Модель міжмасштабного простору кривизни для представлення форми геометричних об'єктів
Пропонується модель міжмасштабного простору кривизни для представлення форми контурів геометричних об'єктів, яка на відміну від відомих моделей масштабно-просторового представлення відображає опуклі ознаки форми в масштабному просторі за рахунок введення міжмасштабного проектування, що дозволяє значно розширити клас об'єктів, розпізнаваних за допомогою масштабно-просторових представлень.

G.A. Kobzar
The model of interscale curvature space is offered for representation of geometric object shapes, which unlike the known models of scale-space representation reflects convex features of shape in scale-space due to introduction of interscale projection that allows extending class of objects recognizable by scale-space representations considerably.

Статья поступила в редакцию 30.01.2008.